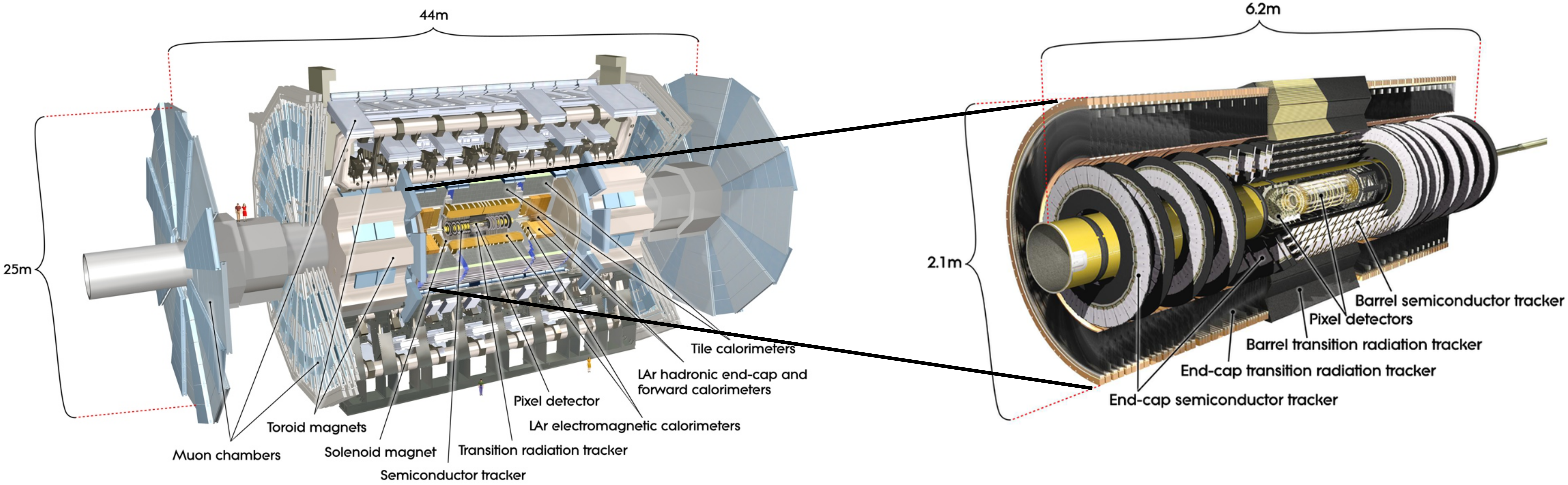
Faster simulated track reconstruction in the ATLAS Fast Chain

Will Leight, for the ATLAS Collaboration
CHEP 2023, Norfolk, VA

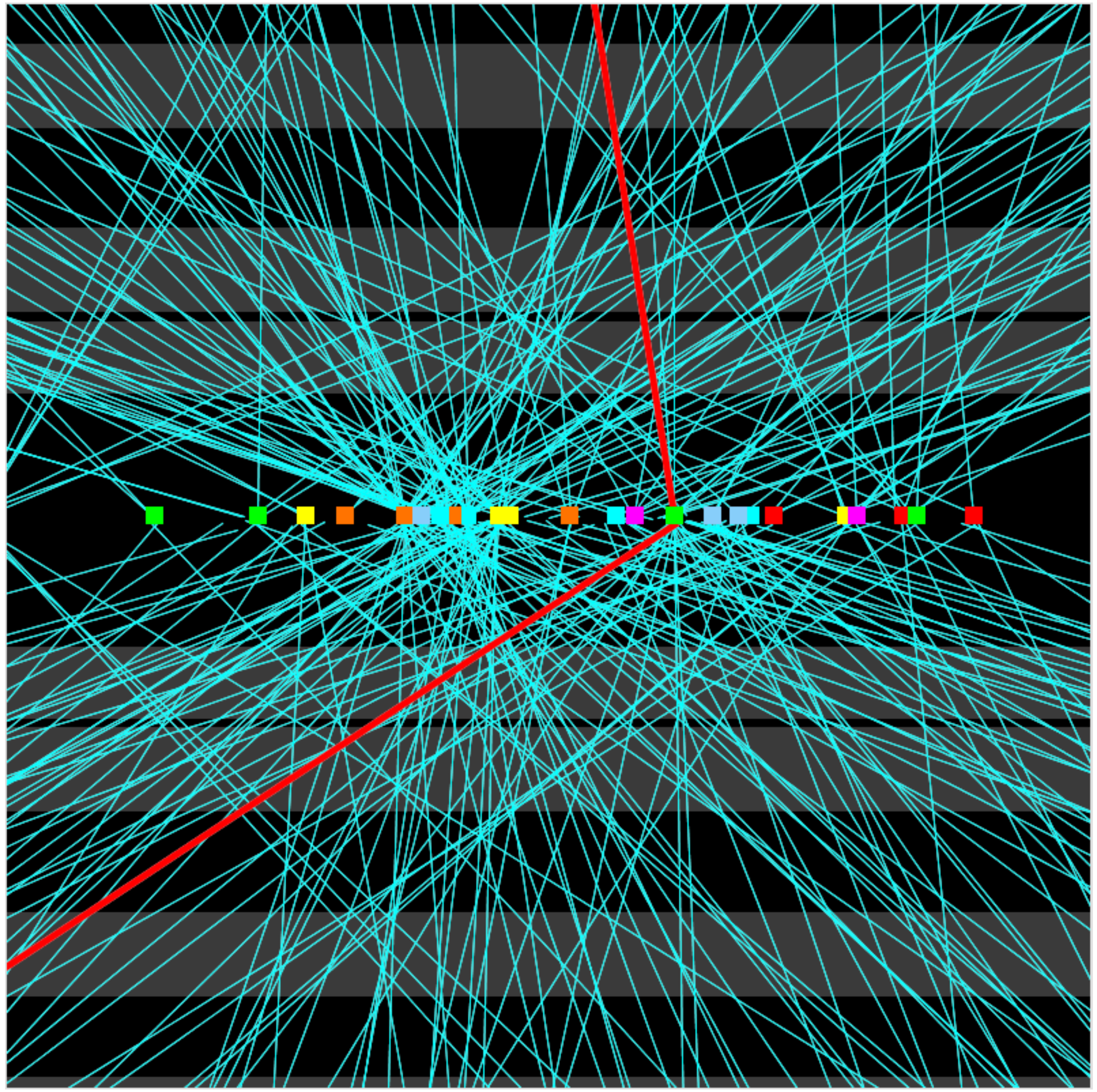
The ATLAS Inner Detector



- Track reconstruction in ATLAS happens in the ID

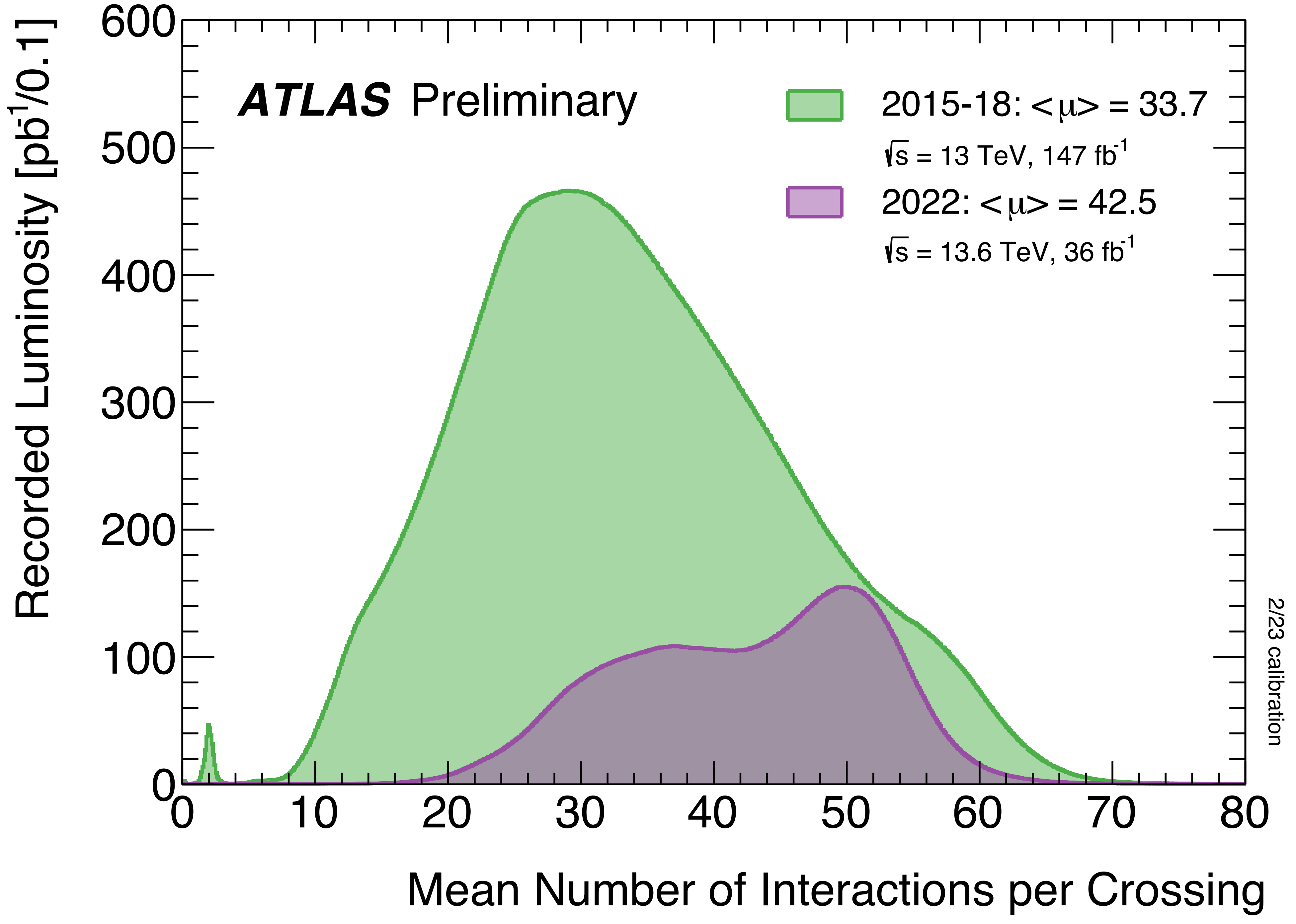
- Multiple layers of silicon detectors, plus transition radiation tracker
  - 2 T solenoidal magnetic field

Pile-Up at the LHC



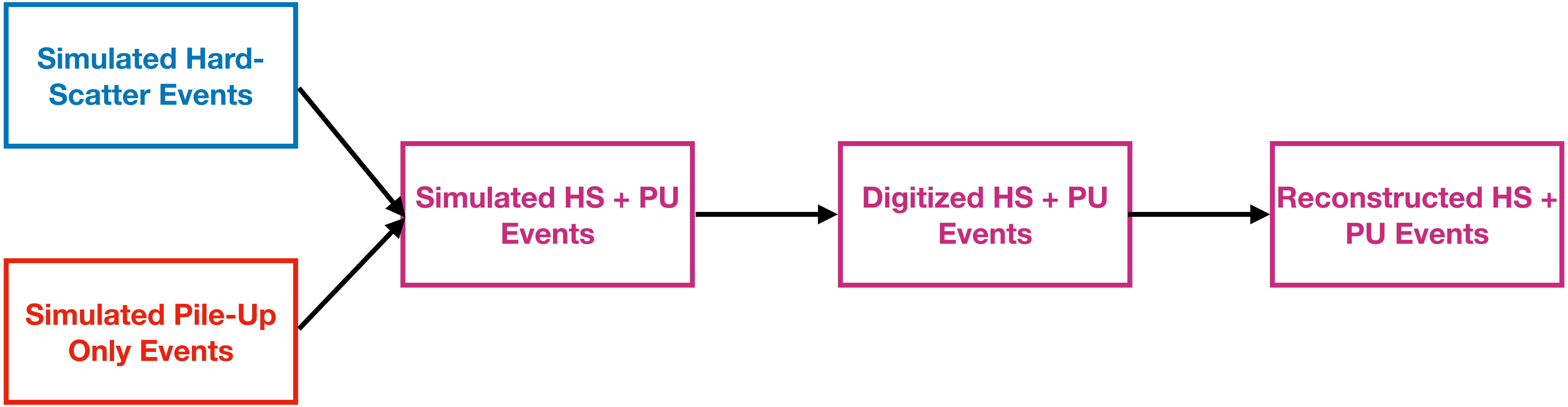
- Every bunch crossing at the LHC comes with many pp collisions
- Only a few will be high-energy collisions
  - These will have the hard-scatter (HS) processes that we are interested in
- That leaves dozens of low-energy pile-up (PU) collisions
  - And their many associated tracks

Pile-Up at the LHC



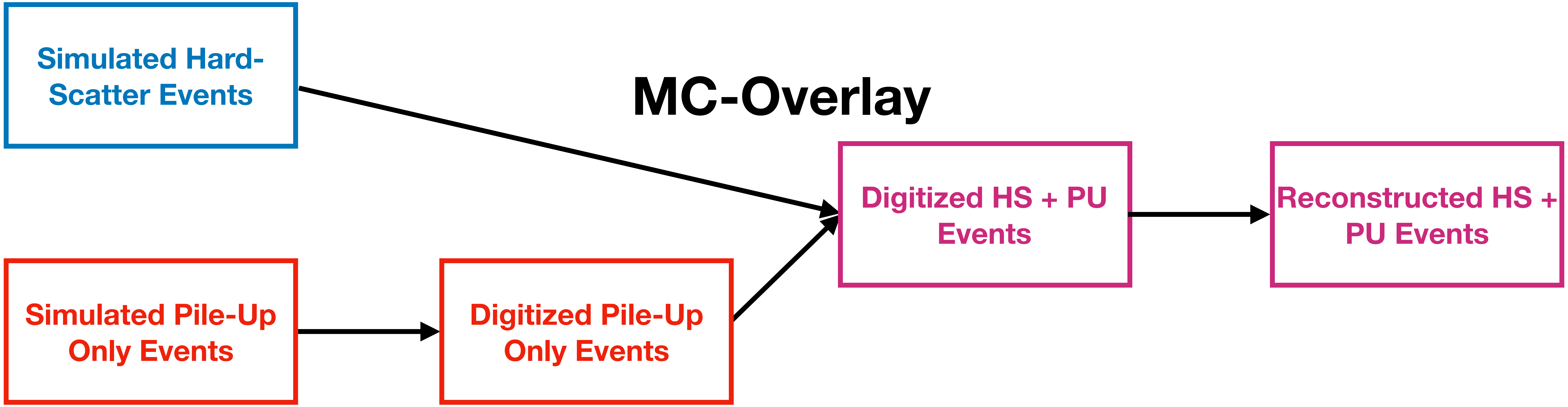
- So far in Run-3, pile-up peaks at ~50 vertices per crossing
  - Significant increase from Run-2
  - For HL-LHC, 200 expected
- Problem for simulating events
  - Simulated events must include pile-up vertices
  - Don't want to spend most simulation time on pile-up vertices
- Solution: simulate pileup separately and reuse it

Simulation With Pile-Up



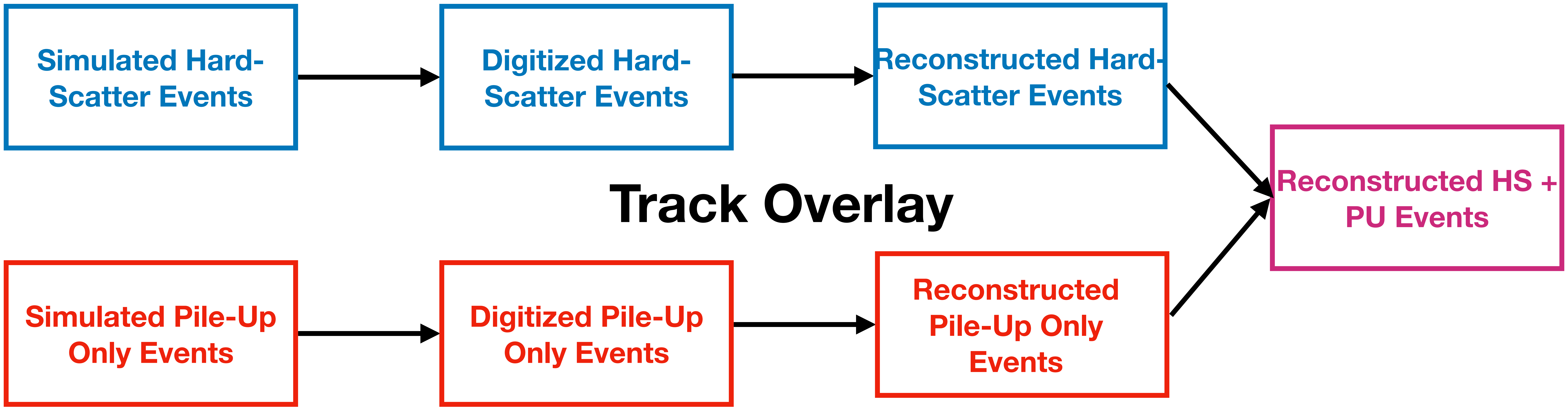
- Pile-up events can be simulated ahead of time and re-used
- Simplest option: merge with the hard-scatter event after simulation

Simulation With Pile-Up



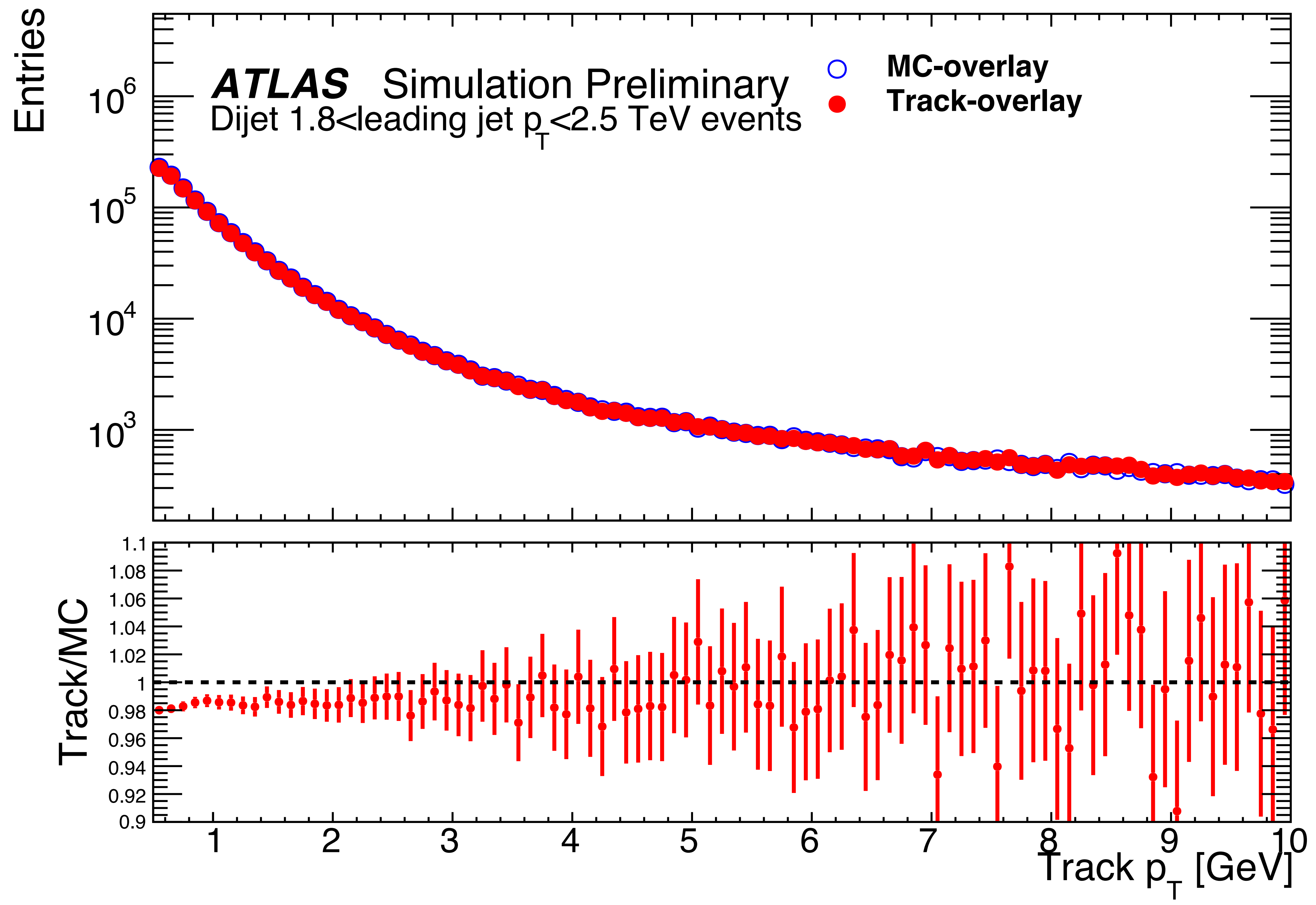
- Pile-up events can be simulated ahead of time and re-used
- Current approach: digitize the pile-up and then merge with the HS event during digitization
  - Comput Softw Big Sci 6, 3 (2022)

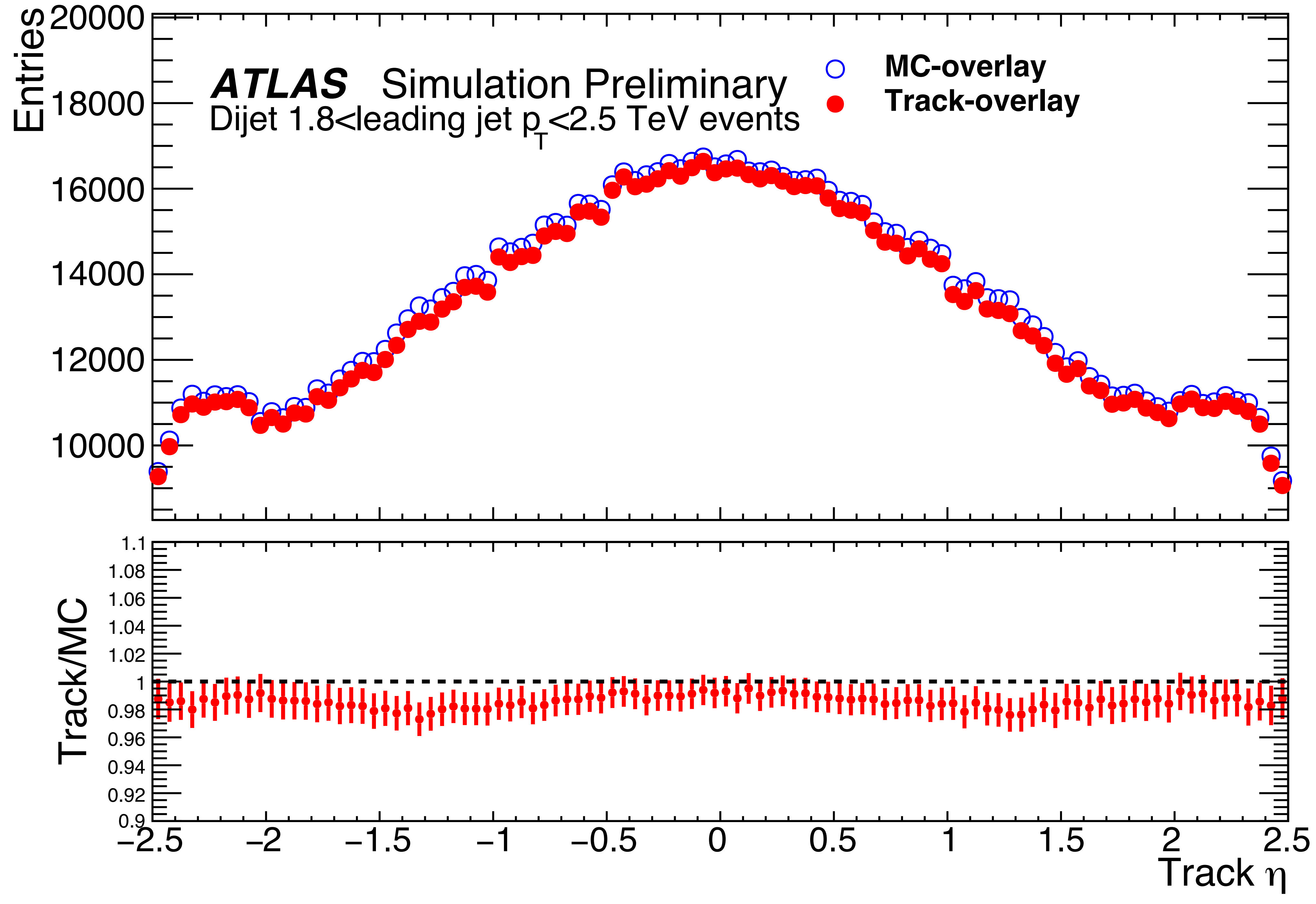
Simulation With Pile-Up



- Pile-up events can be simulated ahead of time and re-used
- Possible new approach: reconstruct both PU and HS separately and combine
  - Only for Inner Detector (ID) tracking
- Good approximation as long as HS tracks don't pick up PU hits

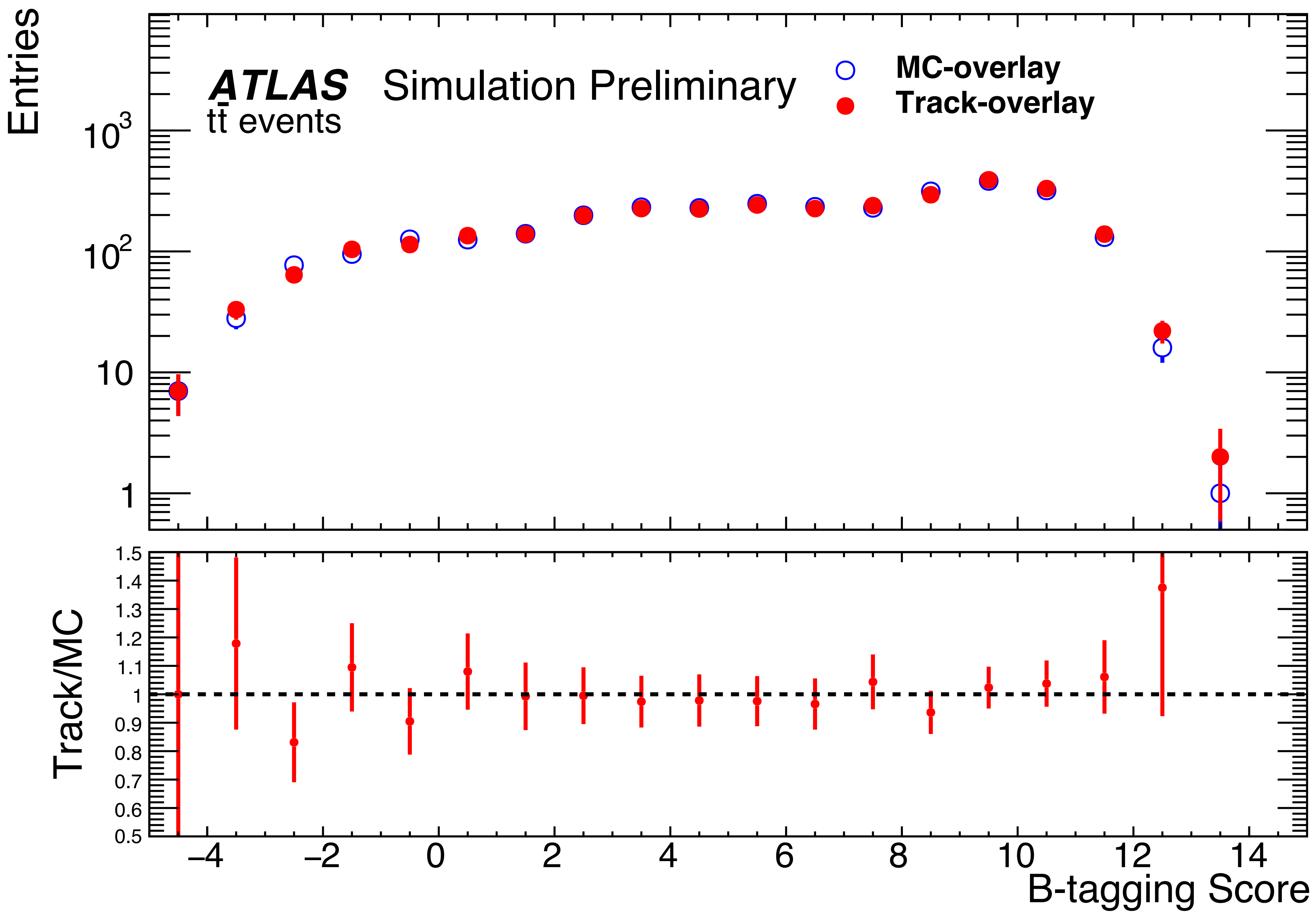
Track Validation

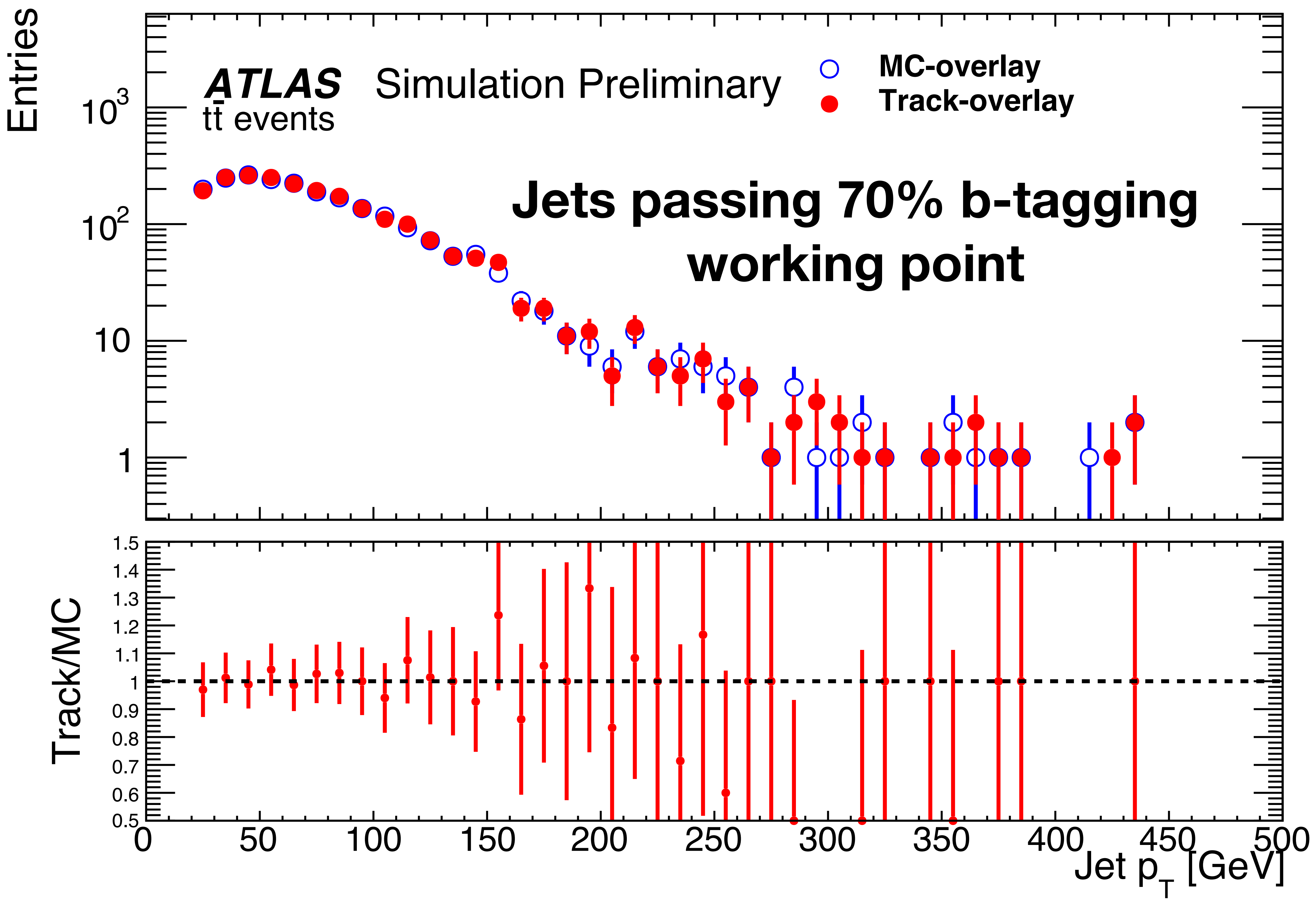




- Good agreement in main track parameters
- Track overlay reconstructs slightly fewer tracks at higher η and lower p_T
  - Fewer fakes when HS and PU hits are not allowed to mix

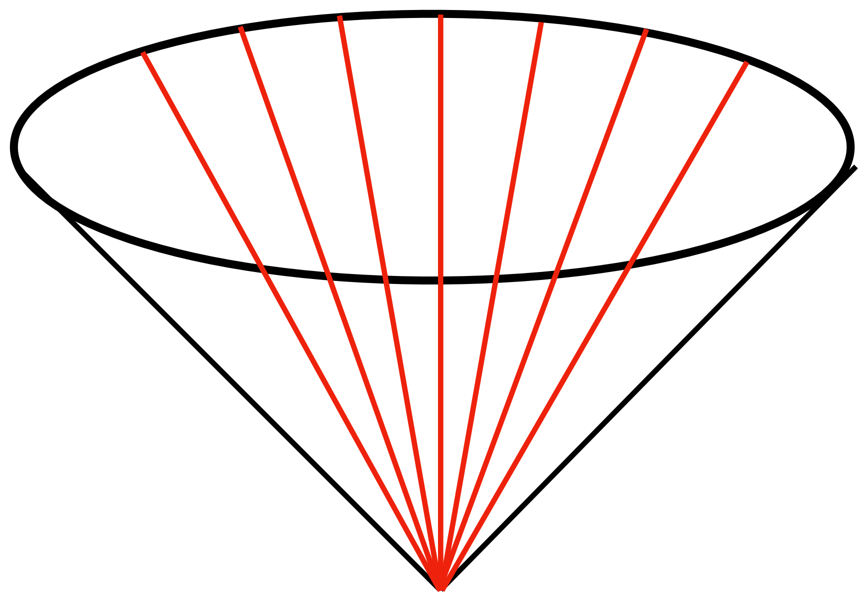
B-jet Validation



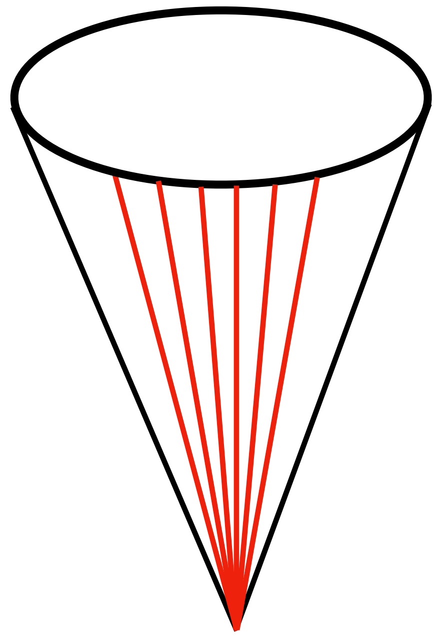


- Part of reconstruction most directly affected by ID tracking changes
- Validation performed using ttbar samples
- No significant disagreement between MC overlay and track overlay

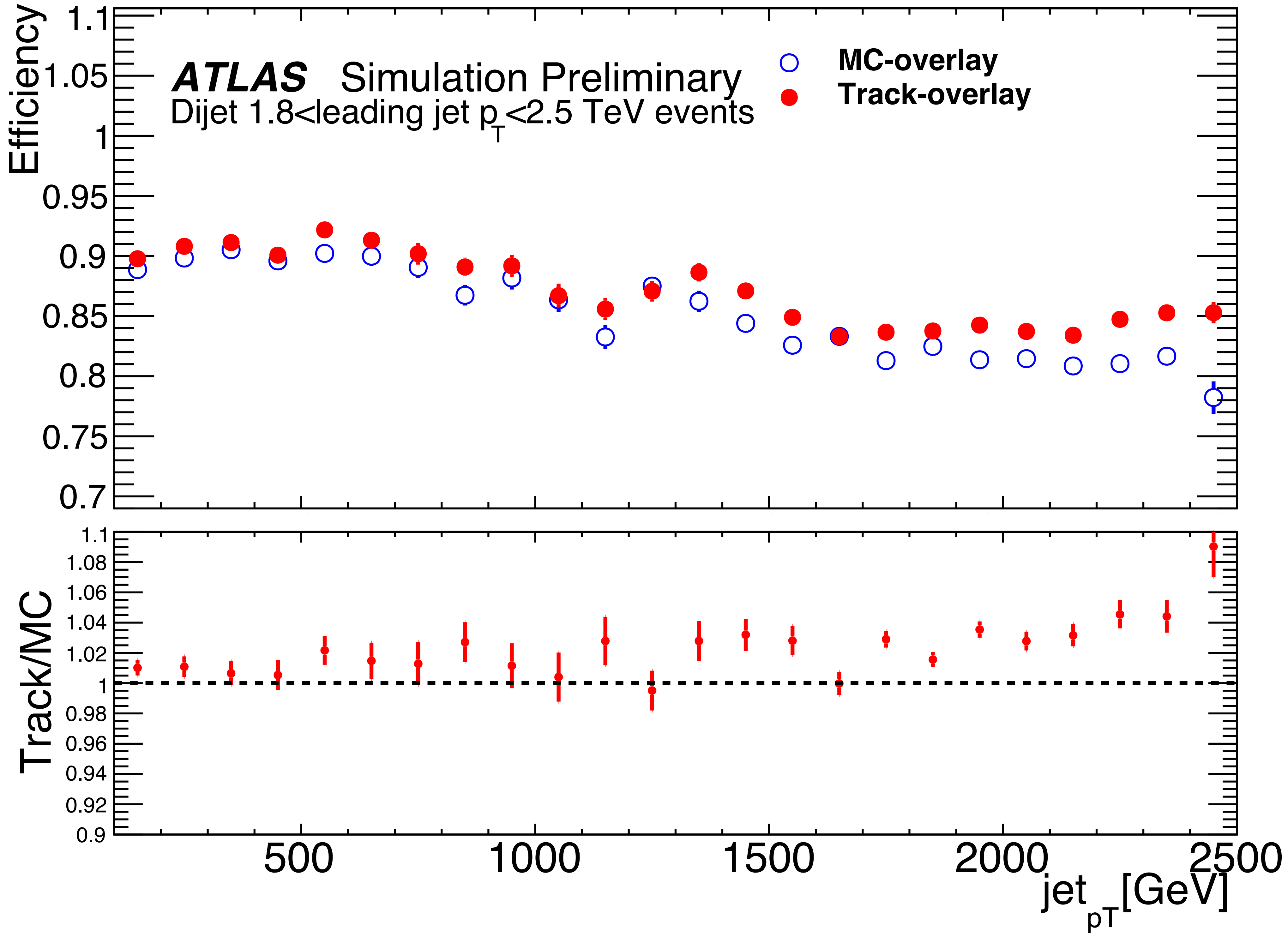
Track-in-Jet Validation



low-p_T jet, low track density

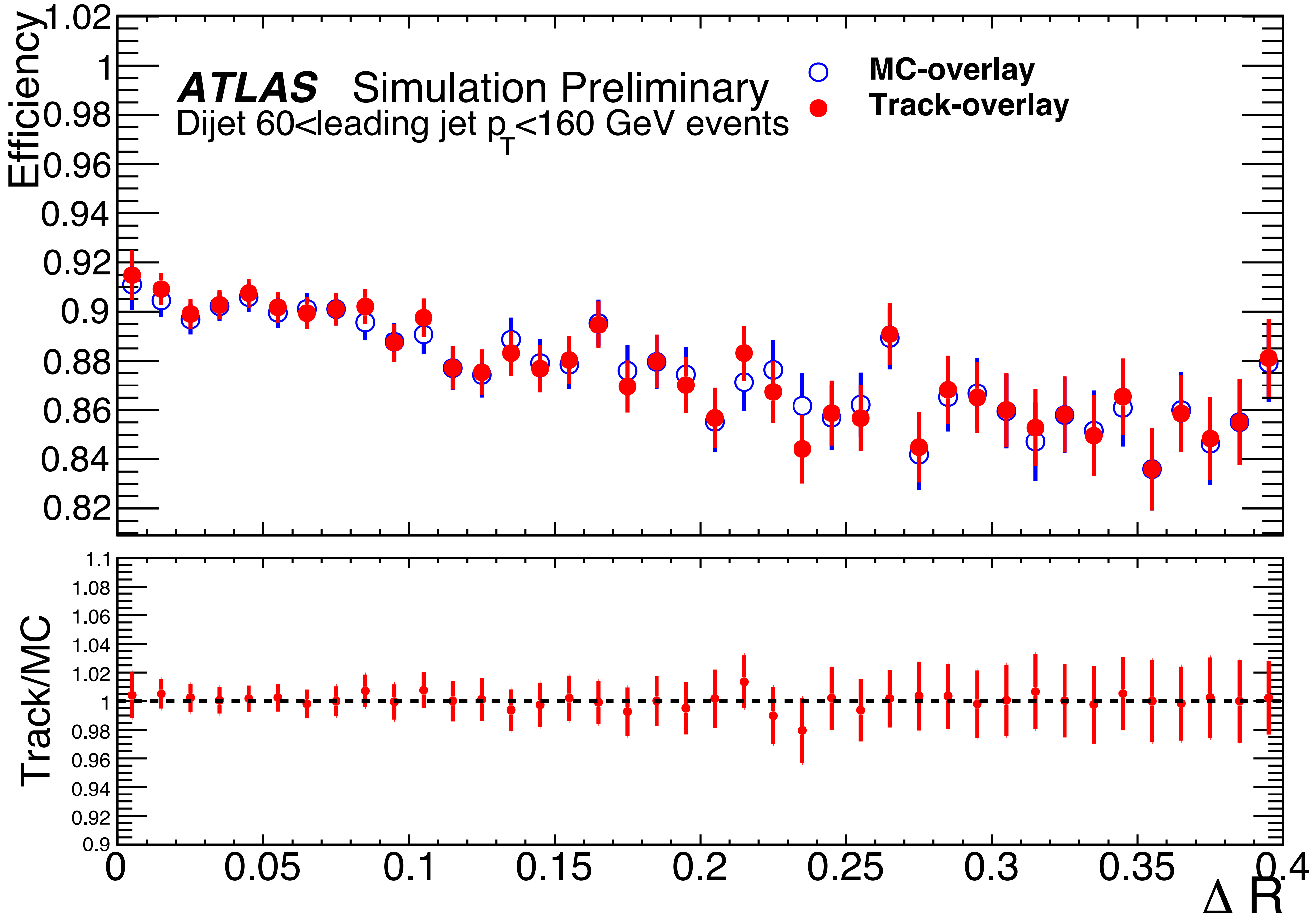


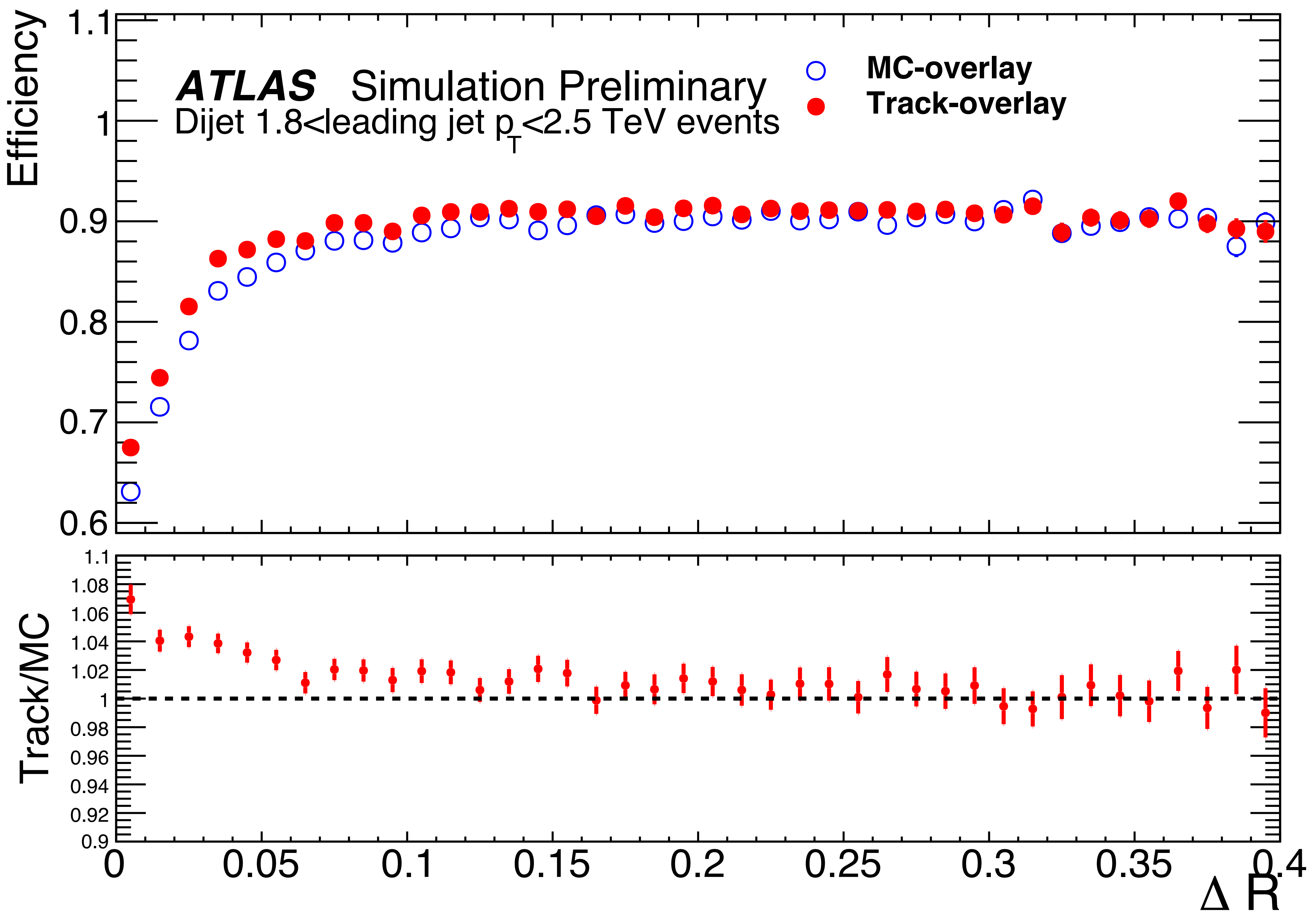
high-p_T jet, high track density



- As expected, more significant differences seen in HS tracks in high-p_T jets
  - Higher-p_T jets are more collimated and so have higher hit densities
  - More likely for PU to effect HS track reconstruction

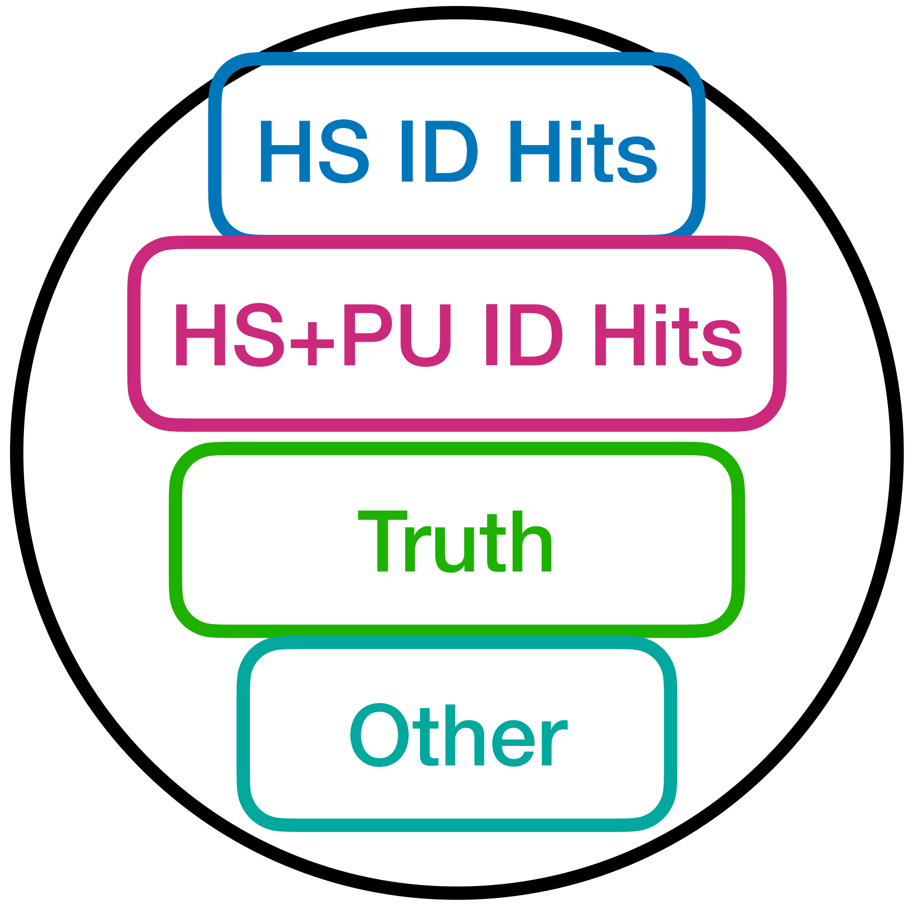
Track-in-Jet Validation





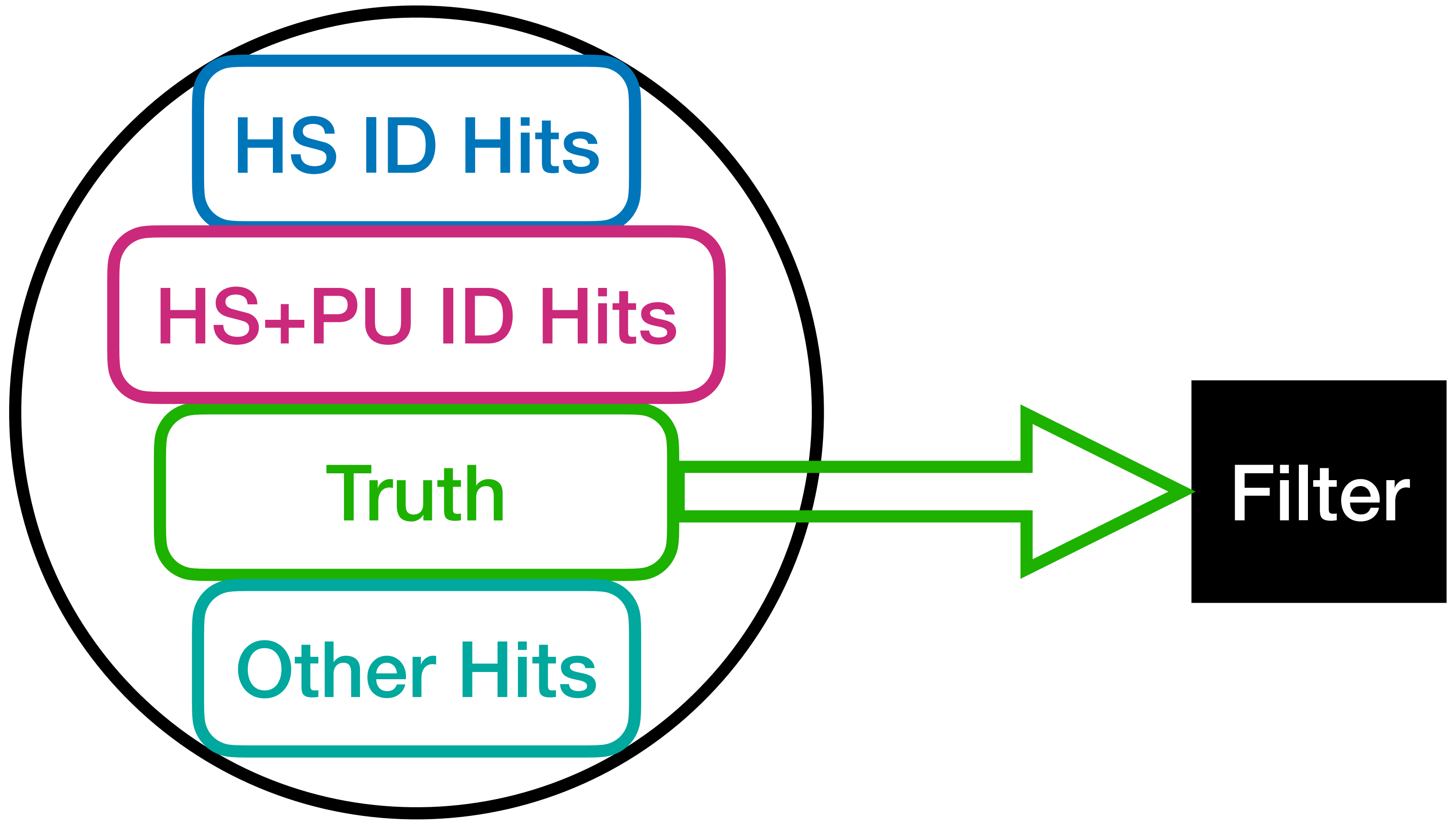
- Difference depends on track location within jets
- In the core of high-p_T jets, track overlay is too efficient
  - Hit density is high enough that reducing it has an observable effect

Overlay Choice



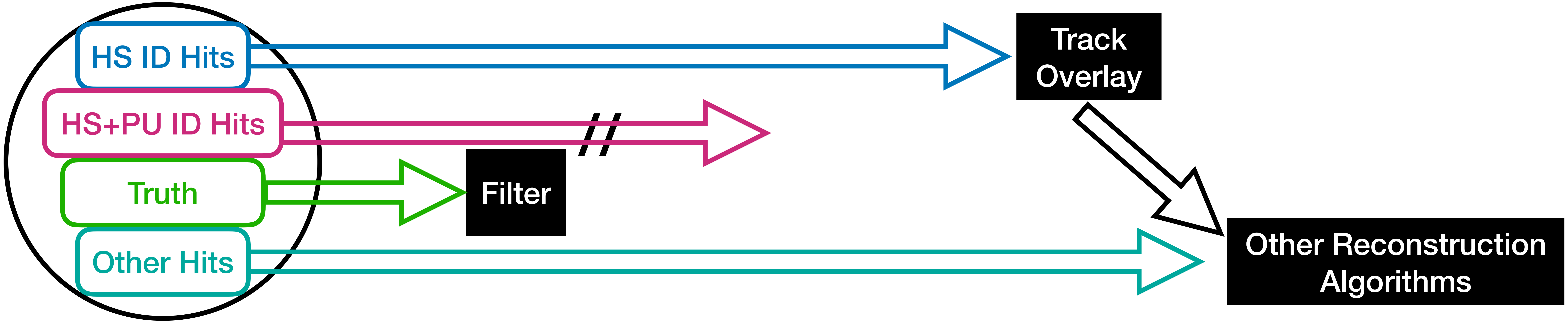
- Track overlay may not be suitable for all events

Overlay Choice



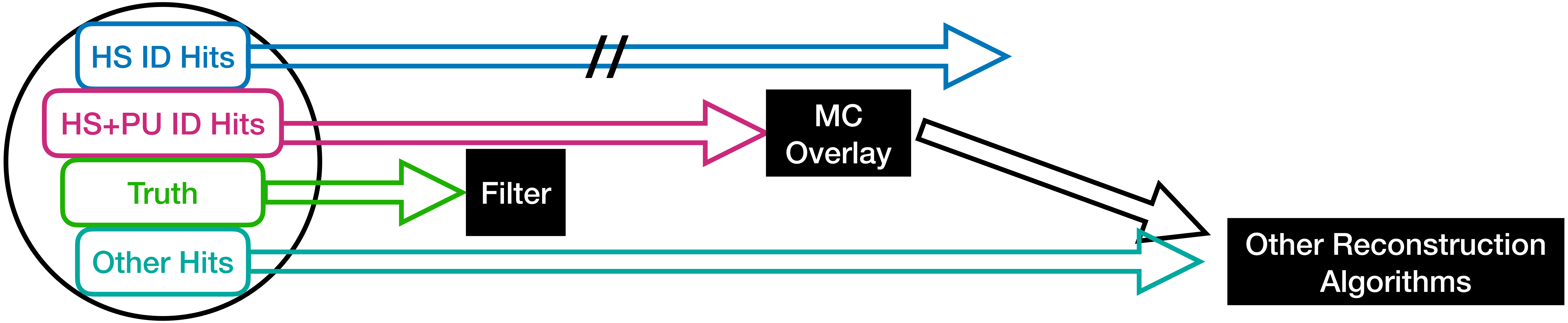
- Track overlay may not be suitable for all events
- Decide event-by-event which events to use it for
  - Train an NN based on truth information

Overlay Choice



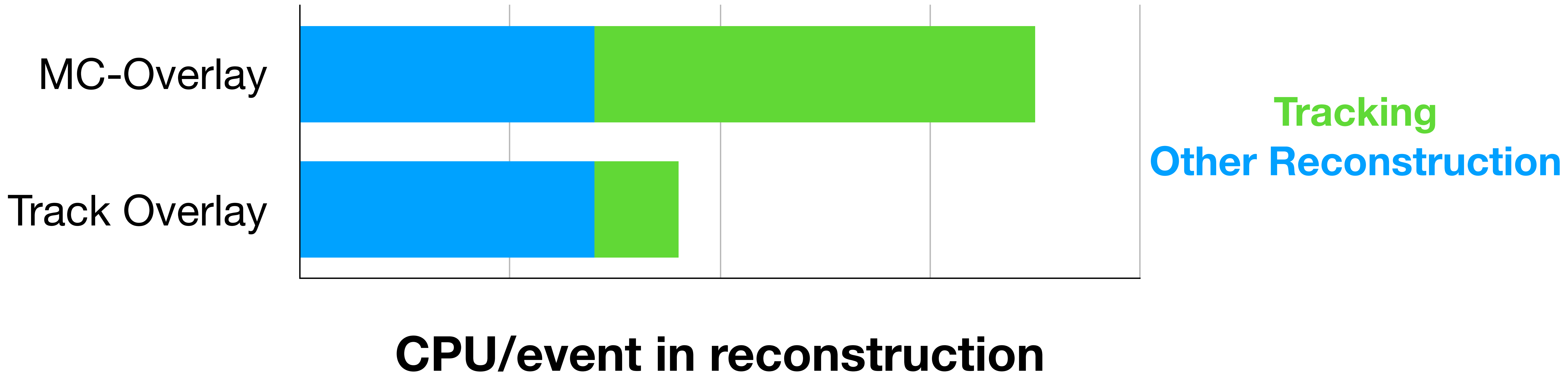
- Track overlay may not be suitable for all events
- Decide event-by-event which events to use it for
  - Train an NN based on truth information
  - Assign events based on the NN output

Overlay Choice



- Track overlay may not be suitable for all events
- Decide event-by-event which events to use it for
  - Train an NN based on truth information
  - Assign events based on the NN output
- Probabilistic assignment of events
  - To ensure a smooth transition between regimes

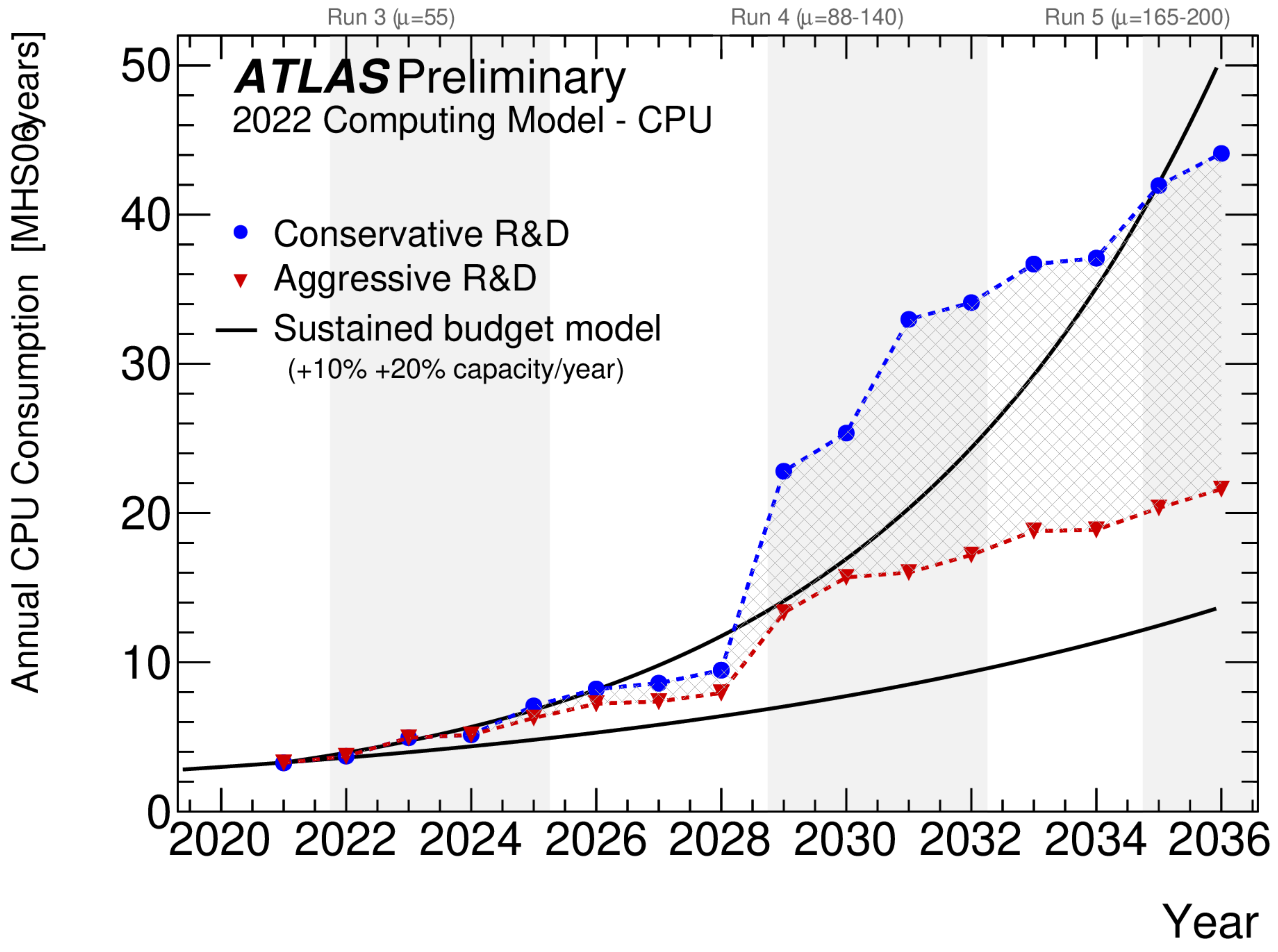
Track Overlay In Run-3

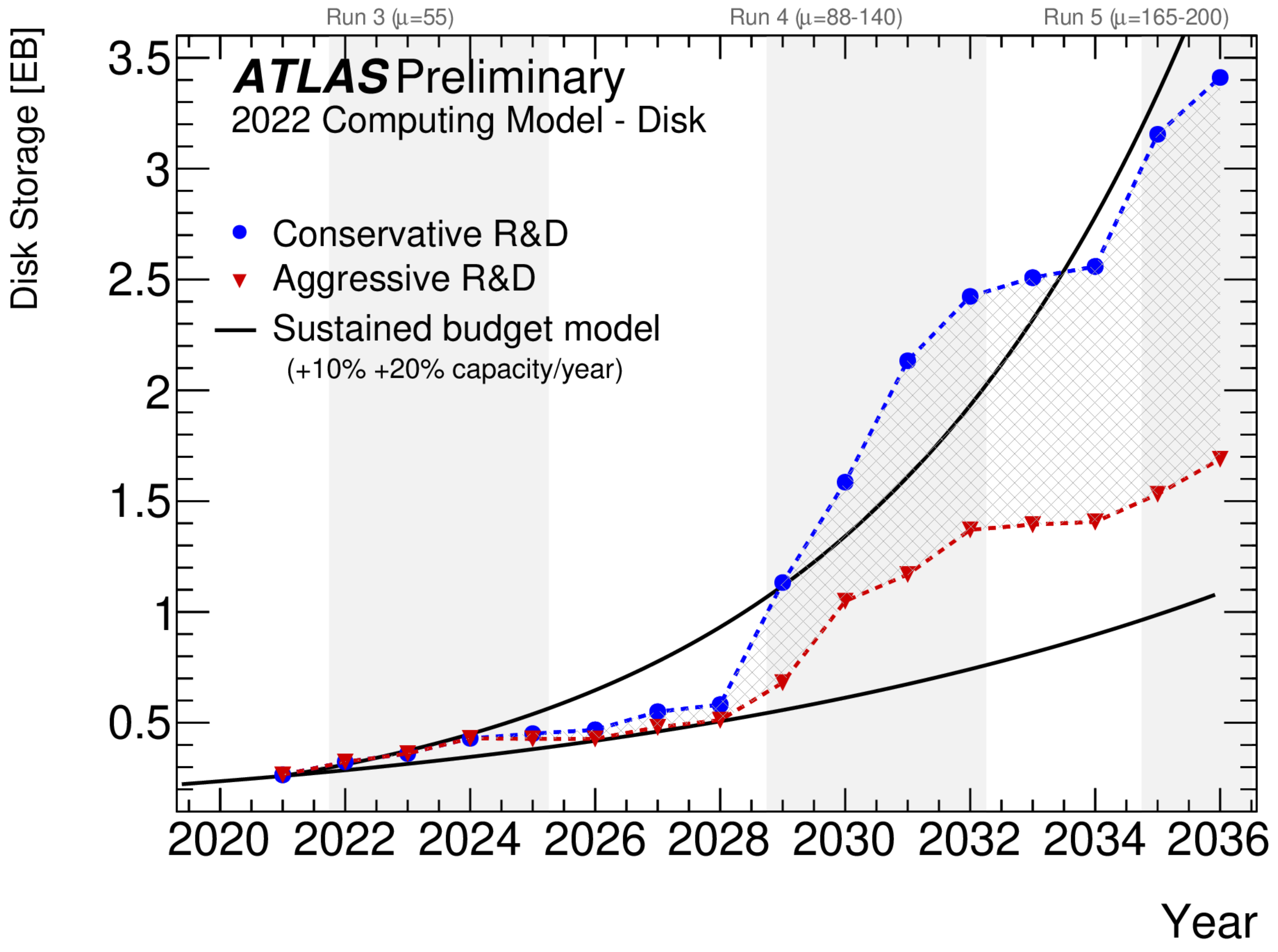


- More than half of event reconstruction time is spent on ID tracking
  - 60% at μ=50

- Track overlay can reduce this by a significant fraction

Computing Challenges for HL-LHC





- HL-LHC will impose considerable demands on CPU and disk usage
- Track overlay is one tool for overcoming these challenges

CPU Reduction for HL-LHC

$\langle\mu\rangle$	primary tracking		unconventional signatures		calorimeter, muon spectr.		combined recon.		monitoring		total recon.	
140	124	(35)	-	(25)	157	(85)	51	(35)	70	(35)	402	(215)
200	214	(50)	-	(30)	176	(95)	94	(70)	100	(50)	584	(295)

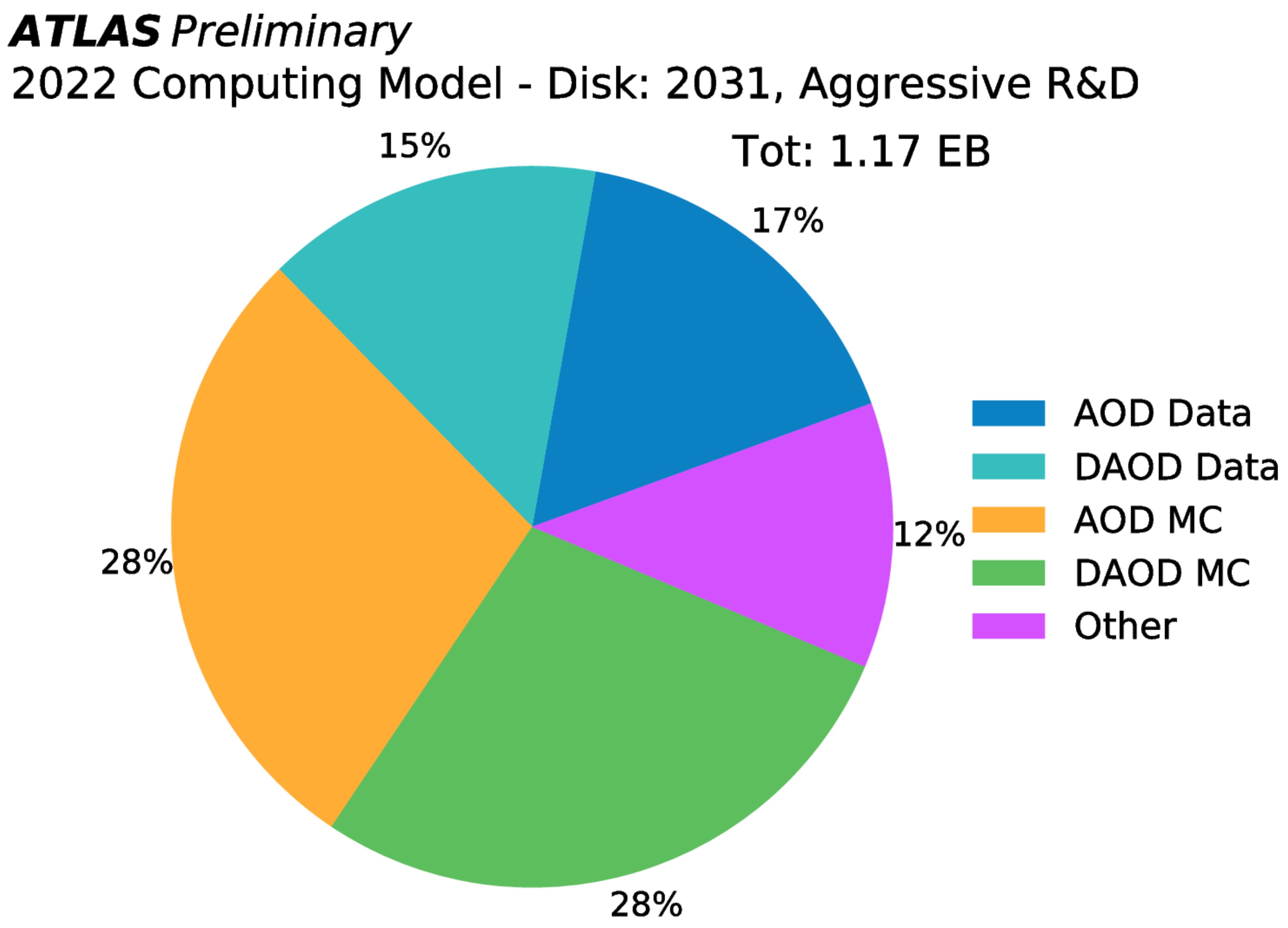
Current Software **Fast Tracking**

- Fast tracking can already achieve considerable reductions
- Track Overlay can remove most of the remaining CPU from tracking
  - Further optimization of muon and calo reco needed to gain a substantial additional speed-up

Disk-use Reduction for HL-LHC

Reconstruction → Intermediate data format: AOD, 500 kB/event → Derivation → Analysis data format: PHYS(LITE), 50 (10) kB/event

- Derivation has to be run several times a year
- Reconstruction speed-up could mean also running reconstruction at that time
- Then no need to save MC AOD
  - Projected 28% disk savings (including data)
  - Also significant saving in tape storage



Summary and Outlook

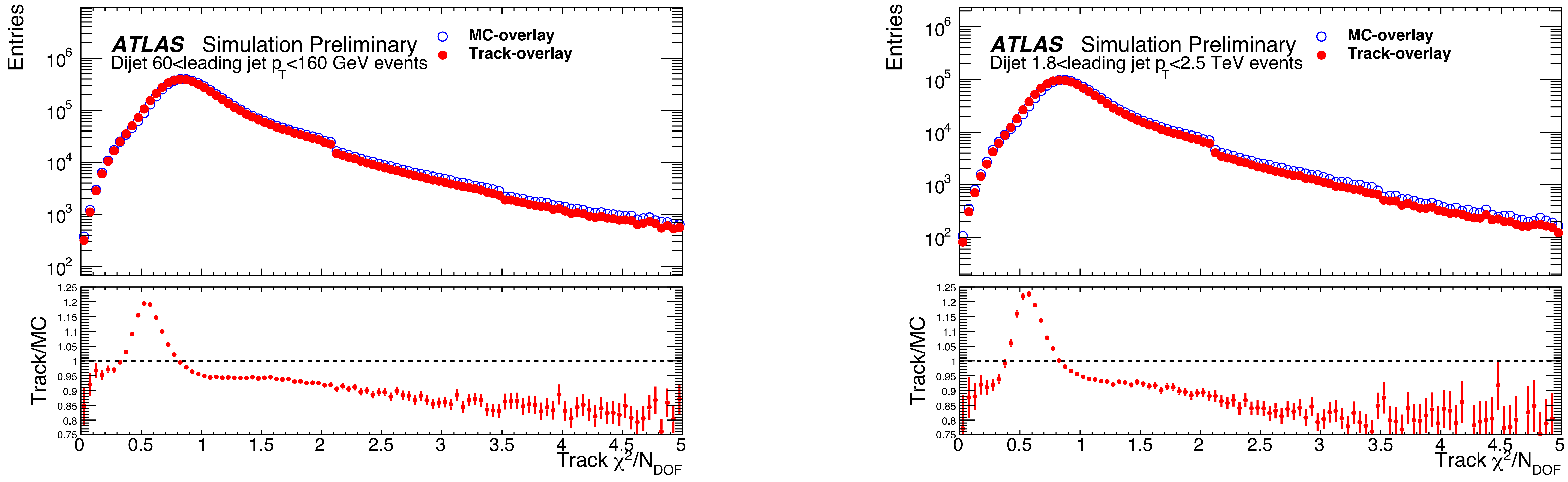
- Reconstructing ID tracks from pile-up ahead of time can speed up simulation
    - One of many necessary improvements to reconstruction
    - Important given that ID tracking takes up the largest fraction of reconstruction time
    - Combined with additional CPU from HPCs or cloud computing, could reduce disk space usage
- Track overlay method now implemented in ATLAS reconstruction
- Good agreement with current method for adding pile-up to events
- Working on implementation of filter to decide which method to use to reconstruct events

Backup

Implementation

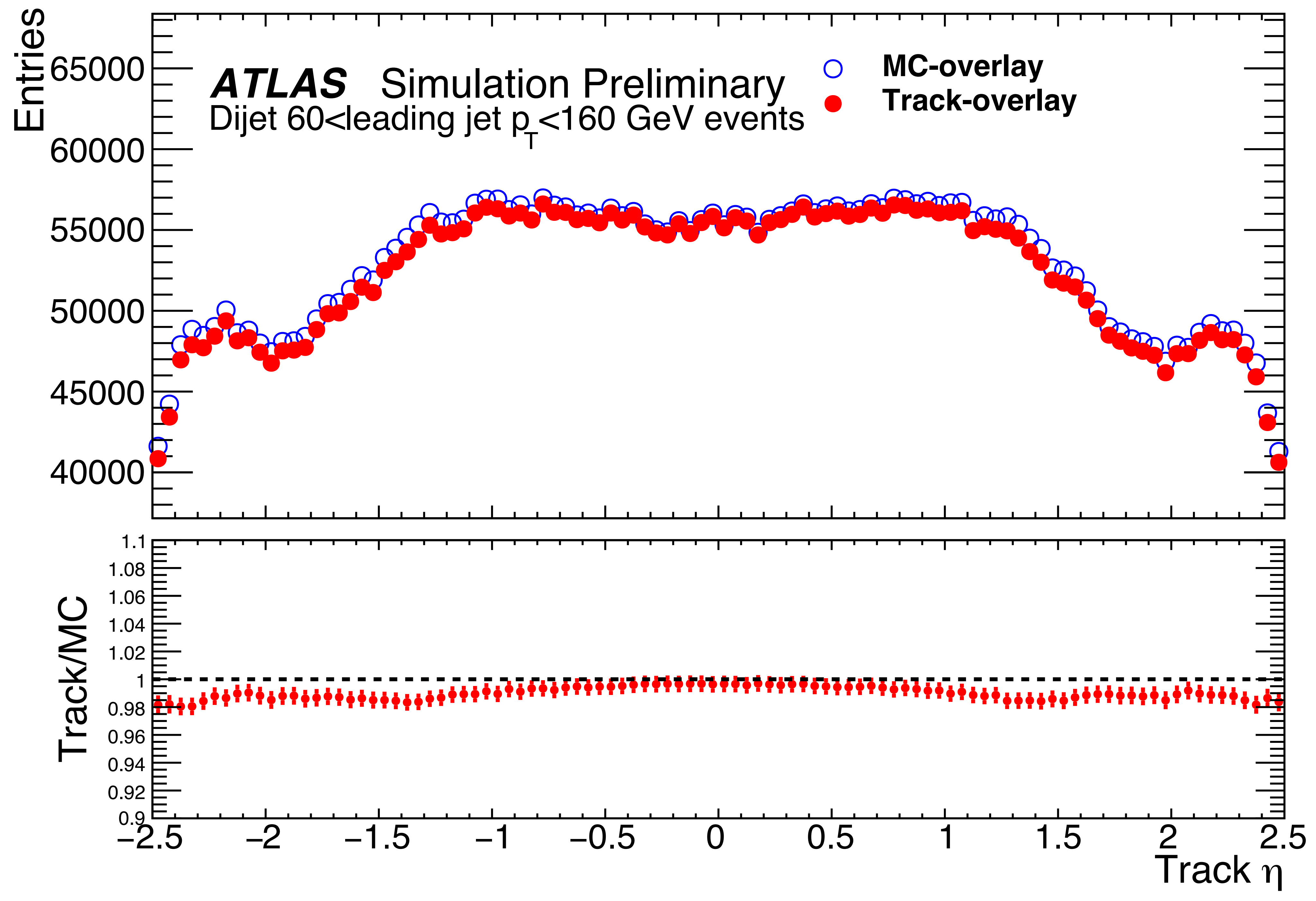
- New setup that runs only ID tracking on simulated pileup file
  - Same tools and parameters as standard reconstruction
  - Generates the same tracking collections
  - Writes out tracks and ID hit clusters
- Pileup tracks and clusters are merged into the final track and cluster collections after HS tracking is run
  - Subsequent reco steps (e.g. muon reco) see the expected full track collections
- Only ID tracking is affected, all other steps in reconstruction proceed as normal

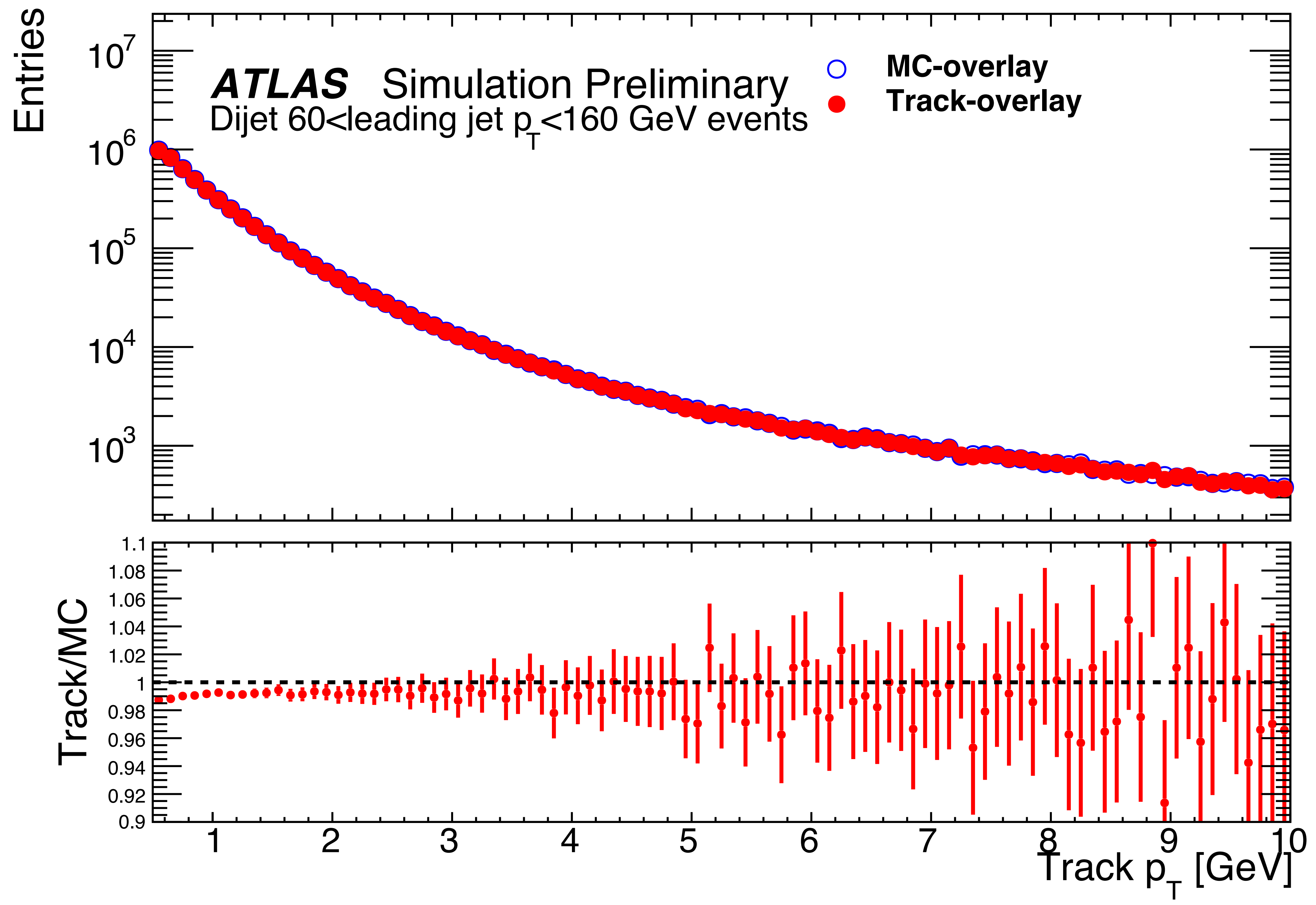
Track Validation



- Validation performed in dijet events
- Track overlay tracks generally have a better χ^2/N_{DOF}
  - Likely due to fewer cases in which the wrong hits are associated to tracks

Track Validation





- Good agreement in track parameters also observed for low-p_T dijet events